

ON THE ASYMPTOTIC MAXIMAL DENSITY OF A SET AVOIDING SOLUTIONS TO LINEAR EQUATIONS MODULO A PRIME

PABLO CANDELA AND OLOF SISASK

ABSTRACT. Given a finite family \mathcal{F} of linear forms with integer coefficients, and a compact abelian group G , an \mathcal{F} -free set in G is a measurable set which does not contain solutions to any equation $L(x) = 0$ for L in \mathcal{F} . We denote by $d_{\mathcal{F}}(G)$ the supremum of $\mu(A)$ over \mathcal{F} -free sets $A \subset G$, where μ is the normalized Haar measure on G . Our main result is that, for any such collection \mathcal{F} of forms in at least three variables, the sequence $d_{\mathcal{F}}(\mathbb{Z}_p)$ converges to $d_{\mathcal{F}}(\mathbb{R}/\mathbb{Z})$ as $p \rightarrow \infty$ over primes. This answers an analogue for \mathbb{Z}_p of a question that Ruzsa raised about sets of integers.

1. INTRODUCTION

Much work in arithmetic combinatorics concerns the maximal density that a subset of a finite abelian group can have if it does not contain a non-trivial solution to a given linear equation. Examples include the study of sum-free sets, where the equation to be avoided is $x_1 + x_2 - x_3 = 0$, and the improvement of bounds for Roth's theorem, which concerns the equation $x_1 - 2x_2 + x_3 = 0$.

A natural question about these maximal densities is whether they exhibit some particular asymptotic behaviour as the groups get larger in a family of groups, and a typical family to consider in this context is that of the groups \mathbb{Z}_p of residues modulo a prime. One may thus ask whether the maximal density of a subset of \mathbb{Z}_p avoiding solutions to some linear equation converges as $p \rightarrow \infty$. The analogous question for subsets of the integers $\{1, \dots, N\}$ was raised by Ruzsa [9, Problem 2.3], who conjectured that the corresponding limit exists for any linear equation. The main result in this paper implies that for the groups \mathbb{Z}_p the maximal density does indeed converge. The result itself is more general, and to state it precisely we shall use the following notation.

Definition 1.1. Let $L(x) = c_1x_1 + \dots + c_tx_t$ be a linear form with non-zero integer coefficients. We say that a subset A of an abelian group G is *L-free* if there is no t -tuple $x = (x_1, \dots, x_t) \in A^t$ such that $L(x) = 0$. For a family \mathcal{F} of linear forms we say that A is *\mathcal{F} -free* if A is L -free for every $L \in \mathcal{F}$.

For a given family \mathcal{F} of linear forms, we define

$$d_{\mathcal{F}}(\mathbb{Z}_p) = \max\{|A|/p : A \subset \mathbb{Z}_p \text{ and } A \text{ is } \mathcal{F}\text{-free}\}.$$

Our main result implies firstly that for any finite family \mathcal{F} of forms in at least 3 variables, the maximal density $d_{\mathcal{F}}(\mathbb{Z}_p)$ converges as $p \rightarrow \infty$ over primes. This is analogous to a

2010 *Mathematics Subject Classification.* Primary 11B30; Secondary 43A25.

Key words and phrases. Sets free from solutions to linear equations, Fourier analysis, removal lemma.

Both authors are EPSRC postdoctoral fellows and gratefully acknowledge the support of the EPSRC.

result of Croot [1] establishing the convergence of the minimal normalised-count of three-term arithmetic progressions in subsets of \mathbb{Z}_p of some fixed density, and indeed a variant of Croot's method will be an essential part of our argument. Before stating the result in full, let us note that convergence can fail in our result if we do not restrict to prime moduli, as was also the case in [1]. Indeed, sum-free sets in \mathbb{Z}_p are easily shown to have maximal density converging to $1/3$ as $p \rightarrow \infty$ over primes (using the Cauchy-Davenport inequality [12]), but in \mathbb{Z}_{2p} the odd residues form a sum-free set of density $1/2$.

The convergence of $d_{\mathcal{F}}(\mathbb{Z}_p)$ leads to the problem of determining the limit. To address this it is potentially helpful to have a fixed group on which the limit can be analyzed. We show that the circle is one possible such group.

Definition 1.2. For a family \mathcal{F} of linear forms and a compact abelian group G with normalized Haar measure μ , we define

$$d_{\mathcal{F}}(G) = \sup\{\mu(A) : A \subset G \text{ is measurable and } \mathcal{F}\text{-free}\}.$$

We can now state our main result.

Theorem 1.3. *Let \mathcal{F} be a finite family of linear forms, each in at least three variables. Then $d_{\mathcal{F}}(\mathbb{Z}_p) \rightarrow d_{\mathcal{F}}(\mathbb{T})$ as $p \rightarrow \infty$ over primes.*

The methods in this paper do not allow to extend this result to families containing forms in two variables, essentially because such forms do not fall under the purview of Fourier analysis; in the language of [5], equations in two variables are of unbounded complexity, while equations in at least three variables are of complexity 1. (Note however that for a single non-trivial form L in two variables it is easy to show that $d_L(\mathbb{Z}_p) \rightarrow 1/2$.)

It is an important fact that if L is a translation-invariant form, that is a form $c_1x_1 + \dots + c_tx_t$ with $c_1 + \dots + c_t = 0$, then $d_L(\mathbb{Z}_p) \rightarrow 0$ as $p \rightarrow \infty$, even if we allow an L -free set to contain certain trivial solutions (constant solutions (x, \dots, x) , for instance); this follows from Roth's method, as observed in a very similar context in [8, Theorem 1.3]. Therefore the limit in Theorem 1.3 is also 0 for any family \mathcal{F} containing a translation-invariant form. On the other hand, if every member of \mathcal{F} is not translation-invariant (to be concise let us say *non-invariant*), then the limit will be positive.

Proposition 1.4. *Let \mathcal{F} be a finite family of non-invariant linear forms, each in at least three variables. Then $d_{\mathcal{F}}(\mathbb{T}) = \lim_{p \rightarrow \infty} d_{\mathcal{F}}(\mathbb{Z}_p) > 0$.*

This situation is thus analogous to that of the maximal edge-densities of graphs with forbidden subgraphs: if one of the forbidden subgraphs is bipartite then the maximal edge-density of a graph on n vertices that does not contain the forbidden subgraphs tends to 0 as $n \rightarrow \infty$; otherwise it tends to a positive limit determined by the chromatic numbers of the forbidden subgraphs.

It would be very interesting to have a simple asymptotic formula for $d_{\mathcal{F}}(\mathbb{Z}_p)$ in terms of the coefficients of the linear forms, if indeed such a formula exists. While Theorem 1.3 does not give such a formula, it does show that it suffices to analyze families of linear equations on \mathbb{T} in order to find one.

The proof of Theorem 1.3 and the remainder of this paper are laid out as follows. In Section 2 we gather tools that will enable us to use Fourier analysis to tackle the

problem. In Section 3 we review and extend a result from [10], itself based on work of Croot [1], that allows us to transfer a solution-free set in \mathbb{Z}_p of density α to an almost-solution-free function on \mathbb{T} with average α . Section 4 then shows how one may use this function to produce an almost-solution-free set of density close to α ; we then use a removal lemma on \mathbb{T} , proved in Section 5, to obtain a truly solution-free subset of \mathbb{T} of density close to α . In Section 6 we combine these results to prove Theorem 1.3, and we also prove Proposition 1.4.

The transference result in Section 3 concerns general compact abelian groups and we shall therefore work in this general setting until the end of that section. From Section 4 onwards we shall focus on the groups \mathbb{Z}_p and \mathbb{T} .

2. PROPERTIES OF S_L

One of the main tools used in this paper is a multilinear operator associated with a given linear form L . This operator, which we denote S_L , is well known in arithmetic combinatorics in the setting of finite groups. In this section we define this operator and establish some of its main properties.

In a finite abelian group G , a set A is L -free in the sense of Definition 1.1 if and only if $|A^t \cap \ker L| = 0$, where t is the number of variables in L and $\ker L$ denotes the subgroup $\{x \in G^t : L(x) = 0\}$ of the direct product G^t . The quantity $|A^t \cap \ker L|/|\ker L|$ is the normalized count of solutions to $L(x) = 0$ inside A . This notion has a natural generalization to any compact abelian group G , using the normalized Haar measure on $\ker L$.

Definition 2.1. Let G be a compact abelian group, let L be a linear form in t variables, and let μ_L denote the Haar measure on the closed subgroup $\ker L$ of G^t satisfying $\mu_L(\ker L) = 1$. Then for any t measurable sets A_1, A_2, \dots, A_t we define the *solution measure* $S_L(A_1, \dots, A_t) = \mu_L((A_1 \times \dots \times A_t) \cap \ker L)$. When $A_1 = A_2 = \dots = A_t = A$ we write more concisely $S_L(A)$.

The operator S_L is the result of extending this definition from sets to functions.

Definition 2.2. Let G be a compact abelian group, let L be a linear form in t variables, and let $f_1, f_2, \dots, f_t : G \rightarrow \mathbb{C}$ be measurable functions. We then define

$$S_L(f_1, \dots, f_t) = \int_{\ker L} f_1 \otimes f_2 \otimes \dots \otimes f_t(x) d\mu_L(x), \quad (2.1)$$

where $f_1 \otimes f_2 \otimes \dots \otimes f_t(x) := f_1(x_1)f_2(x_2) \dots f_t(x_t)$ for all $x = (x_1, \dots, x_t) \in \ker L \subset G^t$. When $f_1 = f_2 = \dots = f_t = f$ we write more concisely $S_L(f)$.

Clearly Definitions 2.1 and 2.2 agree on measurable sets A_1, \dots, A_t , that is we have $S_L(A_1, \dots, A_t) = S_L(1_{A_1}, \dots, 1_{A_t})$, where 1_X denotes the indicator function of a set X .

Let us recall the following standard fact, which will be used throughout the paper.

Lemma 2.3. *Let G and H be compact abelian groups and let $\phi : G \rightarrow H$ be a surjective continuous homomorphism. Then ϕ preserves the normalized Haar measures, that is $\mu_H = \mu_G \circ \phi^{-1}$.*

We shall also use the Gowers U^2 norm.

Definition 2.4. Let $\mathcal{L}_2(G)$ denote the Hilbert space of square-integrable complex-valued functions on G . The U^2 norm $\|\cdot\|_{U^2(G)}$ can be defined on $\mathcal{L}_2(G)$ by the formula

$$\|f\|_{U^2(G)}^4 = \int_{G^4} f(x+y) \overline{f(x+y')} \overline{f(x'+y)} f(x'+y') \, dx \, dx' \, dy \, dy'. \quad (2.2)$$

When the group G in question is clear we write more concisely $\|f\|_{U^2}$.

By a simple application of the Cauchy-Schwarz inequality to the right hand side of (2.2), we obtain the bound $\|f\|_{U^2} \leq \|f\|_{\mathcal{L}_2}$.

For an integer n we let $n \cdot X$ denote the image of $X \subset G$ under the continuous map $x \mapsto nx$. We refer to this map as *dilation by n* or *n -dilation*.

The following result is well known in the setting of finite groups and will be used below.

Theorem 2.5. *Let L be a linear form in $t \geq 3$ variables, and let G be a compact abelian group such that each coefficient of L gives a surjective dilation on G . Then for any $f_1, \dots, f_t \in \mathcal{L}_2(G)$ we have*

$$|S_L(f_1, \dots, f_t)| \leq \min_{i \in [t]} \|f_i\|_{U^2} \prod_{j \neq i} \|f_j\|_{\mathcal{L}_2}. \quad (2.3)$$

Here $[t]$ denotes the set of integers $\{1, 2, \dots, t\}$.

Proof. Writing $L(x) = c_1 x_1 + \dots + c_t x_t$, by assumption we have $c_i \cdot G = G$ for each $i \in [t]$. We shall prove the upper bound in (2.3) just for $i = t$, which is sufficient by symmetry.

By the assumption on c_j -dilations, the map $G^{t-1} \rightarrow \ker L$,

$$(y_1, \dots, y_{t-1}) \mapsto (c_t y_1, c_t y_2, \dots, c_t y_{t-1}, -c_1 y_1 - c_2 y_2 - \dots - c_{t-1} y_{t-1})$$

is surjective, so by Lemma 2.3 we have

$$S_L(f_1, \dots, f_t) = \int_{G^{t-1}} f_1(c_t y_1) \cdots f_{t-1}(c_t y_{t-1}) f_t(-c_1 y_1 - \dots - c_{t-1} y_{t-1}) \, dy_1 \cdots dy_{t-1}. \quad (2.4)$$

We can use this expression to prove the desired upper bound by an argument which is standard in the setting of finite abelian groups, consisting in two applications of the Cauchy-Schwarz inequality. We include the details in the present setting for completeness. First we apply Fubini's theorem and the Cauchy-Schwarz inequality over variables y_2, \dots, y_{t-1} to obtain

$$\begin{aligned} |S_L(f_1, \dots, f_t)|^2 &\leq \int_{G^{t-2}} |f_2(c_t y_2) \cdots f_{t-1}(c_t y_{t-1})|^2 \, dy_2 \cdots dy_{t-1} \\ &\quad \cdot \int_{G^{t-2}} \left| \int_G f_1(c_t y_1) f_t(-c_1 y_1 - \dots - c_{t-1} y_{t-1}) \, dy_1 \right|^2 \, dy_2 \cdots dy_{t-1}. \end{aligned}$$

Applying Lemma 2.3 to c_t -dilation and $(y_2, \dots, y_{t-1}) \mapsto x = -c_2 y_2 - c_3 y_3 - \dots - c_{t-1} y_{t-1}$, the right hand side above is found to equal

$$\|f_2\|_{\mathcal{L}_2}^2 \cdots \|f_{t-1}\|_{\mathcal{L}_2}^2 \int_G \left| \int_G f_1(c_t y) f_t(x - c_1 y) dy \right|^2 dx.$$

By Fubini's theorem the integral here equals

$$\int_{G^2} f_1(c_t y) \overline{f_1(c_t y')} \left(\int_G f_t(x - c_1 y) \overline{f_t(x - c_1 y')} dx \right) dy dy'.$$

Applying Cauchy-Schwarz over (y, y') , Lemma 2.3 for dilations by c_1 and by c_t , and Fubini's theorem again, we find that this integral is at most $\|f_1\|_{\mathcal{L}_2}^2$ times

$$\left(\int_{G^4} f_t(x + y) \overline{f_t(x + y')} \overline{f_t(x' + y)} f_t(x' + y') dx dx' dy dy' \right)^{1/2} = \|f_t\|_{U^2}^2. \quad \square$$

Note that, since the \mathcal{L}_2 norm dominates the U^2 norm, (2.3) implies immediately

$$|S_L(f_1, \dots, f_t)| \leq \prod_i \|f_i\|_{\mathcal{L}_2}. \quad (2.5)$$

We now turn to Fourier-analytic aspects of S_L . Let us first settle on some notation. For a compact abelian group G and any real $r \geq 1$ we denote by $\mathcal{L}_r(G)$ the Banach space of complex-valued functions on G with integrable r th power. The implicit measure here is the normalized Haar measure on G . The Pontryagin dual \widehat{G} of G is a discrete group, so by $\mathcal{L}_r(\widehat{G})$ we denote the analogous Banach space but with the implicit measure being the counting measure on \widehat{G} , which assigns value 1 to each singleton.

We shall use the Fourier transform on $\mathcal{L}_2(G)$ (or Plancherel transform). By the compactness of G we have $\mathcal{L}_2(G) \subset \mathcal{L}_1(G)$. If $f \in \mathcal{L}_1(G)$ then the Fourier transform \widehat{f} is defined for $\gamma \in \widehat{G}$ by $\widehat{f}(\gamma) = \int_G f(x) \overline{\gamma(x)} d\mu_G(x)$. If \widehat{f} is also in $\mathcal{L}_1(\widehat{G})$ then we have the Fourier inversion formula

$$f(x) = \int_{\widehat{G}} \widehat{f}(\gamma) \gamma(x) d\mu_{\widehat{G}}(\gamma). \quad (2.6)$$

Plancherel's theorem gives us that for any $f \in \mathcal{L}_2(G)$, $\widehat{f} \in \mathcal{L}_2(\widehat{G})$ and $\|f\|_{\mathcal{L}_2(G)} = \|\widehat{f}\|_{\mathcal{L}_2(\widehat{G})}$.

The following expression of S_L in terms of the Fourier transforms \widehat{f}_i will be used in the next section.

Proposition 2.6. *Let $L = c_1 x_1 + \dots + c_t x_t$ be a linear form and let G be a compact abelian group such that $c_i \cdot G = G$ for every i . Then for any $f_1, \dots, f_t \in \mathcal{L}_2(G)$ we have*

$$S_L(f_1, \dots, f_t) = \int_{\widehat{G}} \widehat{f}_1(\gamma^{c_1}) \cdots \widehat{f}_t(\gamma^{c_t}) d\mu_{\widehat{G}}(\gamma). \quad (2.7)$$

Proof. Let $H = \ker L$. First we prove (2.7) for f_1, \dots, f_t having Fourier transforms in $\mathcal{L}_1(\widehat{G})$. In this case each f_i is continuous (by (2.6)), and it follows that the function F defined on G^t/H by $F(x + H) = \int_H f(x + y) d\mu_H(y)$ is continuous on G^t/H . Note that $S_L(f_1, \dots, f_t) = F(0)$. Let us consider the Fourier coefficients $\widehat{F}(\chi)$, $\chi \in \widehat{G^t/H} \cong H^\perp$.

On one hand, it follows from $c_i \cdot G = G$ that $\gamma \mapsto \chi = \gamma \circ L$ is an isomorphism $\widehat{G} \rightarrow H^\perp$. On the other hand, using a standard formula for integration on quotient groups [7, 2.7.3 (2)] one checks that $\widehat{F}(\gamma \circ L) = \int_{G^t} f(x) \overline{\gamma \circ L(x)} d\mu_{G^t}(x) = \widehat{f}_1(\gamma^{c_1}) \cdots \widehat{f}_t(\gamma^{c_t})$. Therefore, provided \widehat{F} is in $\mathcal{L}_1(H^\perp)$, we can apply Fourier inversion to conclude that

$$S_L(f_1, \dots, f_t) = F(0) = \int_{H^\perp} \widehat{F}(\chi) d\mu_{H^\perp}(\chi) = \int_{\widehat{G}} \widehat{f}_1(\gamma^{c_1}) \cdots \widehat{f}_t(\gamma^{c_t}) d\mu_{\widehat{G}}(\gamma).$$

The function $\gamma \mapsto \widehat{f}_1(\gamma^{c_1}) \cdots \widehat{f}_t(\gamma^{c_t})$ is shown to be indeed in $\mathcal{L}_1(\widehat{G})$ using Cauchy-Schwarz, Plancherel's theorem, and the fact that $\gamma \mapsto \gamma^{c_i}$ is injective (i.e. that $c_i \cdot G = G$).

Now let $f_i \in \mathcal{L}_2(G)$, so $\widehat{f}_i \in \mathcal{L}_2(\widehat{G})$ with $\|\widehat{f}_i\|_{\mathcal{L}_2(\widehat{G})} = \|f_i\|_{\mathcal{L}_2(G)}$. For each i we have a sequence $\widehat{g}_{i,n}$ in $\mathcal{L}_1(\widehat{G})$ with $\widehat{g}_{i,n} \rightarrow \widehat{f}_i$ in the $\mathcal{L}_2(\widehat{G})$ norm, whence also $g_{i,n} \rightarrow f_i$ in $\mathcal{L}_2(G)$. One then uses multilinearity of S_L , (2.5), and Cauchy-Schwarz on $\mathcal{L}_2(\widehat{G})$ to show that

$$\begin{aligned} S_L(f_1, \dots, f_t) &= \lim_{n \rightarrow \infty} S_L(g_{1,n}, g_{2,n}, \dots, g_{t,n}) \\ &= \lim_{n \rightarrow \infty} \int_{\widehat{G}} \widehat{g}_{1,n}(\gamma^{c_1}) \cdots \widehat{g}_{t,n}(\gamma^{c_t}) d\mu_{\widehat{G}}(\gamma) = \int_{\widehat{G}} \widehat{f}_1(\gamma^{c_1}) \cdots \widehat{f}_t(\gamma^{c_t}) d\mu_{\widehat{G}}(\gamma). \end{aligned}$$

□

We close this section with the observation that for the proof of Theorem 1.3 we can assume that each $L \in \mathcal{F}$ has coprime coefficients. This is justified by the following lemma.

Lemma 2.7. *Let \mathcal{F} be a finite family of linear forms, and let \mathcal{F}' be obtained by multiplying each $L \in \mathcal{F}$ by some non-zero integer n_L . Let G be a compact abelian group such that $n_L \cdot G = G$ for every L . Then $d_{\mathcal{F}}(G) = d_{\mathcal{F}'}(G)$.*

Proof. It is clear that any nL -free set is also L -free, so $d_{\mathcal{F}'}(G) \leq d_{\mathcal{F}}(G)$. On the other hand, if A is L -free then $n^{-1}A$ is nL -free for any n , for if $x_i \in n^{-1}A$ and $(nc_1)x_1 + \cdots + (nc_t)x_t = 0$ then $(nx_1, \dots, nx_t) \in A^t$ is a solution. If in addition $n \cdot G = G$ then $\mu(n^{-1}A) = \mu(A)$. These properties imply easily that if A is \mathcal{F} -free then $m^{-1}A$ is \mathcal{F}' -free, where $m = \prod_{L \in \mathcal{F}} n_L$, and so $d_{\mathcal{F}}(G) \leq d_{\mathcal{F}'}(G)$. □

3. TRANSFERENCE

In proving Theorem 1.3 we shall need to move sets between the groups \mathbb{Z}_p and \mathbb{T} . A result essentially allowing us to do so was established in Chapter 4 of [10] by extending ideas of Croot [1], though for simplicity it was assumed there that all linear forms considered had at least one coefficient equal to 1. In this section we shall review this result and also show how to eliminate the assumption on the coefficients, thus obtaining Proposition 3.4 and, as our main application, Corollary 3.8.

Central to the results we are about to discuss is the notion of Freiman isomorphism.

Definition 3.1. Let $k \geq 2$ and let $A \subset G$, $B \subset H$ be subsets of two abelian groups. We call a function $\varphi : A \rightarrow B$ a *Freiman k -isomorphism* if it is a bijection and

$$a_1 + \cdots + a_k = a_{k+1} + \cdots + a_{2k} \iff \varphi(a_1) + \cdots + \varphi(a_k) = \varphi(a_{k+1}) + \cdots + \varphi(a_{2k})$$

for all $a_i \in A$.

Thus Freiman k -isomorphisms, or just k -isomorphisms for short, preserve additive relations of length at most k . The main result of [10, Chapter 4] was that if one can find k -isomorphisms between small subsets of the duals of G and H , then one can model functions on one group by functions on the other, in a particular sense. That sense uses the following notion of *admissibility* of a linear form.

Definition 3.2. Let G be a compact abelian group and let $L = c_1x_1 + \cdots + c_tx_t$ be a linear form in at least three variables with coprime coefficients. We say that L is *G -admissible* if $c_i \cdot G = G$ for each coefficient c_i . Since the c_i are coprime there are integers n_i such that $n_1c_1 + \cdots + n_tc_t = 1$; we call the minimum value of $|n_1| + \cdots + |n_t|$ over such integers the *multiplier-height* of L and we denote this quantity by $h(L)$. We shall say that L is *k -admissible* if

$$k \geq \max\{h(L), |c_1|, |c_2|, \dots, |c_t|\}. \quad (3.1)$$

We give the obvious meaning to (G, k) -admissibility and, if H is another compact abelian group, to (G, H, k) -admissibility.

One more definition is needed in order to state the transference result.

Definition 3.3. Let G and H be abelian groups. We say that H *can Freiman (n, k) -model G* if for any subset A of G of size $|A| \leq n$ there exists a Freiman k -isomorphism from A to a subset of H .

Proposition 3.4. Let $\epsilon > 0$ and h be a positive integer. Suppose G and H are two compact abelian groups such that, for some $k \geq k_0(\epsilon)$ and all $n \leq C(\epsilon)^h$, \widehat{H} can Freiman (n, k) -model \widehat{G} . Let $f : G \rightarrow [0, 1]$ be a measurable function with $\int_G f = \alpha$. Then there is a continuous function $g : H \rightarrow [0, 1]$ with $\int_H g = \alpha$ such that

$$|S_L(f) - S_L(g)| \leq t\epsilon\alpha^{t-2}$$

for any (G, H, k) -admissible linear form L in $t \geq 3$ variables with multiplier-height at most h .

The main difference between this proposition and [10, Proposition 4.2.8] lies in the parameter h : since the latter result only dealt with forms where at least one coefficient was equal to 1 one could always take $h = 1$. The purpose of the rest of this section is to indicate the proof of [10, Proposition 4.2.8] and show how one may adapt it to take the parameter h into account.

Step 1: regularize f . The proof begins with a modification of a standard procedure: one replaces f by a function $f' : G \rightarrow [0, 1]$ with Fourier support $R \subset \widehat{G}$ of bounded size, with R containing the identity $1_{\widehat{G}}$ and with $R = R^{-1}$, so that

$$f'(x) = \sum_{\gamma \in R} \widehat{f'}(\gamma) \gamma(x)$$

for all $x \in G$ and $|R| \leq C(\epsilon)$. One can do this in such a way that $\int_G f' = \int_G f$ and $|S_L(f) - S_L(f')| \leq t\epsilon\alpha^{t-2}/2$ for any G -admissible linear form L in t variables.

Step 2: transfer to H . By the assumption that \widehat{H} can (n, k) -model \widehat{G} one can find a k -isomorphism $\varphi : R \rightarrow R' \subset \widehat{H}$. One then defines $g : H \rightarrow \mathbb{R}$ by

$$g(x) := \sum_{\gamma \in R} \widehat{f}'(\gamma) \varphi(\gamma)(x).$$

The key properties of g used in [10] were that $\int_H g = \int_G f$ and that $S_L(g) = S_L(f')$ for any (G, H, k) -admissible form L that has at least one coefficient equal to 1; we shall therefore need to modify this part of the argument.

Step 3: control the range of g . The function g produced in the previous step does not a priori take values in $[0, 1]$, as we need it to do. However, one can then produce a function $g' : H \rightarrow [0, 1]$ such that $\int_H g' = \int_H g$ and $|S_L(g) - S_L(g')| \leq t\epsilon\alpha^{t-2}/2$ for any H -admissible linear form L (there are several ways to do this; see [10]). This step completes the proof.

Proof of Proposition 3.4. We shall only need to modify Step 2 above. We are thus given a function $f : G \rightarrow \mathbb{R}$ and a symmetric set $R \subset \widehat{G}$ containing the identity $1_{\widehat{G}}$ such that

$$f(x) = \sum_{\gamma \in R} \widehat{f}(\gamma) \gamma(x)$$

for all $x \in G$, and we have that R has bounded size: $|R| \leq C(\epsilon)$. Let $Q \subset \widehat{G}$ be the set $Q := R^h = R \cdot R \cdots R$ where h is the height-parameter supplied to Proposition 3.4. Certainly we have $|Q| \leq C(\epsilon)^h$, and so the modelling hypothesis guarantees the existence of a Freiman k -isomorphism $\varphi : Q \rightarrow Q' \subset \widehat{H}$. By translating we may assume that $\varphi(1_{\widehat{G}}) = 1_{\widehat{H}}$. Note that this means that

$$\gamma_1^{r_1} \cdots \gamma_n^{r_n} = \gamma_1^{s_1} \cdots \gamma_n^{s_n} \iff \varphi(\gamma_1)^{r_1} \cdots \varphi(\gamma_n)^{r_n} = \varphi(\gamma_1)^{s_1} \cdots \varphi(\gamma_n)^{s_n}$$

whenever $\gamma_i \in Q$ and the r_i, s_i are non-negative integers with $\sum r_i \leq k$ and $\sum s_i \leq k$. Note also that $\varphi(\gamma^{-1}) = \varphi(\gamma)^{-1}$ for all $\gamma \in Q$ since $\varphi(1_{\widehat{G}}) = 1_{\widehat{H}}$. We now establish the following lemma, which replaces [10, Lemma 4.6.3].

Lemma 3.5. *Define $g : H \rightarrow \mathbb{R}$ by setting*

$$g(x) := \sum_{\gamma \in R} \widehat{f}(\gamma) \varphi(\gamma)(x) \tag{3.2}$$

for each $x \in H$. Then $\int_H g = \int_G f$, and if L is a (G, H, k) -admissible linear form of multiplier-height at most h then $S_L(g) = S_L(f)$.

Proof. The properties follow from the fact that (3.2) is the Fourier expansion of g and from the Fourier-inversion of S_L provided by Proposition 2.6. Indeed,

$$\widehat{g}(\chi) = \begin{cases} \widehat{f}(\gamma) & \text{if } \chi = \varphi(\gamma) \text{ for some } \gamma \in R, \\ 0 & \text{otherwise.} \end{cases}$$

Hence $\int_H g = \widehat{g}(1_{\widehat{H}}) = \widehat{f}(1_{\widehat{G}}) = \int_G f$. Now let $L = c_1 x_1 + \cdots + c_t x_t$ be a (G, H, k) -admissible form of multiplier-height at most h ; thus we have integers n_1, \dots, n_t such

that $n_1c_1 + \dots + n_tc_t = 1$ and $|n_1| + \dots + |n_t| \leq h$. Proposition 2.6 gives that

$$S_L(f) = \sum_{\substack{\gamma \in \widehat{G} \\ \gamma^{c_i} \in R \forall i}} \widehat{f}(\gamma^{c_1}) \dots \widehat{f}(\gamma^{c_t}) \quad \text{and} \quad S_L(g) = \sum_{\substack{\chi \in \widehat{H} \\ \chi^{c_i} \in \varphi(R) \forall i}} \widehat{g}(\chi^{c_1}) \dots \widehat{g}(\chi^{c_t}).$$

Let us write $\Gamma := \{\gamma \in \widehat{G} : \gamma^{c_i} \in R \text{ for all } i\}$ and $\Psi := \{\chi \in \widehat{H} : \chi^{c_i} \in \varphi(R) \text{ for all } i\}$ for the index sets occurring in these two sums. The result will follow if we can show that φ is a bijection from Γ to Ψ such that $\varphi(\gamma)^{c_i} = \varphi(\gamma^{c_i})$ for all i and $\gamma \in \Gamma$. We certainly have the second property, for if $\gamma \in \Gamma$ then

$$\gamma = \gamma^{n_1c_1 + \dots + n_tc_t} = (\gamma^{c_1})^{n_1} \dots (\gamma^{c_t})^{n_t} \in R^{n_1} \dots R^{n_t} \subset Q,$$

and φ is a k -isomorphism on Q , where $k \geq |c_i|$ for all i . So we just need to establish that $\varphi(\Gamma) = \Psi$. Let us first deal with $\varphi(\Gamma) \subset \Psi$: let $\gamma \in \Gamma$. Then we need to show that $\varphi(\gamma)^{c_i} \in \varphi(R)$ for all i . But this is immediate since $\varphi(\gamma)^{c_i} = \varphi(\gamma^{c_i})$. The opposite inclusion follows in the same way using φ^{-1} since $\Psi \subset \varphi(Q)$, which follows from the fact that if $\chi \in \Psi$ then

$$\chi = (\chi^{c_1})^{n_1} \dots (\chi^{c_t})^{n_t} \in \varphi(R)^{n_1} \dots \varphi(R)^{n_t} = \varphi(R^{n_1} \dots R^{n_t}) \subset \varphi(Q),$$

φ being a Freiman h -isomorphism. □

The rest of the proof of Proposition 3.4 is identical to that of the proof of Proposition 4.2.8 in [10], and so we are done. □

Proposition 3.4 gives us a criterion for transferring functions between two compact abelian groups. The following lemmas show that this criterion allows us to work with the groups \mathbb{Z}_p and \mathbb{T} , since $\widehat{\mathbb{Z}_p} \cong \mathbb{Z}_p$ and $\widehat{\mathbb{T}} \cong \mathbb{Z}$.

Lemma 3.6. *Let $n, k \in \mathbb{N}$ and let $p \geq (2k)^n$ be a prime. Then for any set $A \subseteq \mathbb{Z}_p$ of size n there is a set $B \subset \mathbb{Z}$ that is Freiman k -isomorphic to A . In other words, \mathbb{Z} can Freiman (n, k) -model \mathbb{Z}_p provided $p \geq (2k)^n$.*

This result is standard and follows from an application of Dirichlet's box principle. The proof of the following lemma from [10] is slightly more subtle but still elementary.

Lemma 3.7. *Let n, k be positive integers and let $N \geq (4k)^n$ be an integer. Then for any set $A \subseteq \mathbb{Z}$ of size n there is a set $B \subset \mathbb{Z}_N$ that is Freiman k -isomorphic to A . In other words, \mathbb{Z}_N can Freiman (n, k) -model \mathbb{Z} provided $N \geq (4k)^n$.*

Thus we obtain the following immediate corollary of Proposition 3.4.

Corollary 3.8. *Let $\epsilon > 0$ and h be a positive integer, and let each of G and H be \mathbb{Z}_p or \mathbb{T} , where $p \geq C(\epsilon, h)$ is a prime. Then for any measurable function $f : G \rightarrow [0, 1]$ with $\int_G f = \alpha$ there is a continuous function $g : H \rightarrow [0, 1]$ with $\int_H g = \alpha$ such that*

$$|S_L(f) - S_L(g)| \leq t\epsilon\alpha^{t-2}$$

for any (G, H, k) -admissible linear form L in $t \geq 3$ variables with multiplier-height at most h .

4. FROM FUNCTIONS TO SETS

In our proof of Theorem 1.3 we shall use Proposition 3.4 to obtain a function g with certain properties, and we shall then require a set with similar properties. The existence of such a set will be guaranteed by the following result.

Lemma 4.1. *Let $\epsilon > 0$, let G be \mathbb{T} or \mathbb{Z}_p for p sufficiently large, and let $f : G \rightarrow [0, 1]$ be measurable. Then there exists a measurable set $A \subset G$ such that $|\mu_G(A) - \int_G f| \leq \epsilon$ and $|S_L(A) - S_L(f)| \leq t\epsilon$ for any G -admissible linear form L in $t \geq 3$ variables.*

Note that any linear form is \mathbb{T} -admissible and most forms are \mathbb{Z}_p -admissible. To prove this lemma we shall use Theorem 2.5 and a probabilistic construction that is familiar in the setting of finite abelian groups.

Given two sets A, B let $A \Delta B$ denote their symmetric difference. The following notion of discretization will be used here and in the next section.

Definition 4.2. A set $A \subset \mathbb{T}$ is (δ, n) -measurable if there exists a set B which is the union of dyadic intervals $I_{n,j} := [(j-1)/2^n, j/2^n)$, $j \in [2^n]$, such that $\mu(A \Delta B) < \delta$.

A function $f : \mathbb{T} \rightarrow [0, 1]$ is (δ, n) -measurable if there is a function $g : \mathbb{T} \rightarrow [0, 1]$ that is constant on the intervals $I_{n,j}$ such that $\|f - g\|_{\mathcal{L}_1} < \delta$.

We shall use both of the following equivalent instances of Littlewood's first principle, which are discussed in [11, §2.4].

Lemma 4.3. *Let $A \subset \mathbb{T}$ be measurable. Then for every $\delta > 0$ there exists n such that A is (δ, n) -measurable.*

Lemma 4.4. *Let $f : \mathbb{T} \rightarrow [0, 1]$ be measurable. Then for every $\delta > 0$ there exists n such that f is (δ, n) -measurable.*

While one cannot in general approximate a $[0, 1]$ -valued function by a set in $\mathcal{L}_1(G)$, one can do so in $U^2(G)$ in the sense of the following result. It is this that allows us to establish Lemma 4.1.

Lemma 4.5. *Let $\epsilon > 0$, let G be \mathbb{T} or \mathbb{Z}_p for p sufficiently large, and let $f : G \rightarrow [0, 1]$ be measurable. Then there is a measurable set $A \subset G$ such that $\|f - 1_A\|_{U^2} \leq \epsilon$. Moreover, for $G = \mathbb{T}$ we may take A to be a finite union of dyadic intervals.*

Proof. For $G = \mathbb{Z}_p$ this is a standard result: define a subset A of G randomly by letting $x \in A$ with probability $f(x)$ independently for each $x \in G$. Then by independence the expectation of $\|f - 1_A\|_{U^2}^4$ equals 0, up to an error of magnitude $O(p^{-1})$ due to averaging over degenerate parallelograms (i.e. parallelograms $(x+y, x+y', x'+y, x'+y')$ with at least two vertices being equal). Thus there exists a choice of A for which $\|f - 1_A\|_{U^2} = O(p^{-1/4})$.

Now we adapt this standard argument to deal with the group \mathbb{T} . First we approximate f by a suitable step-function: by Lemma 4.4 there is an integer n and coefficients $\gamma_j \in [0, 1]$ such that the function $g = \sum_{j \in [2^n]} \gamma_j 1_{I_{n,j}}$ satisfies $\|f - g\|_{\mathcal{L}_1} < \epsilon^4$. We thus have $\|f - g\|_{U^2} \leq \|f - g\|_{\mathcal{L}_1}^{1/4} < \epsilon$.

Next we define an appropriate finite probability space. Consider the set Ω of functions of the form $1_A(x) = \sum_{j \in [2^n]} \alpha_j 1_{I_{n,j}}(x)$, where for all j we have $\alpha_j \in \{0, 1\}$. An element in Ω can be identified with an element $\alpha = (\alpha_1, \dots, \alpha_{2^n})$ of $\{0, 1\}^{2^n}$, and we can then define a probability on Ω by declaring the events $\{\alpha_j = 1\}$ to be independent and assigning to $\{\alpha_j = 1\}$ the probability γ_j .

Now we compute the expectation of $\|g - 1_A\|_{U^2}^4$ for a randomly chosen $1_A \in \Omega$, using the following familiar expression for the U^2 norm:

$$\|f\|_{U^2(G)}^4 = \int_{G^3} f(x) \overline{f(x+h)} \overline{f(x+k)} f(x+h+k) dx dh dk.$$

This expression is seen to equal (2.2) by Lemma 2.3. By Fubini's theorem we have

$$\mathbb{E}\|g - 1_A\|_{U^2}^4 = \int_{\mathbb{T}^3} \mathbb{E}(g - 1_A)(x)(g - 1_A)(x+h)(g - 1_A)(x+k)(g - 1_A)(x+h+k) dx dh dk.$$

Fix any values for x, h, k such that no two vertices of the corresponding parallelogram $(x, x+h, x+k, x+h+k)$ lie in the same interval $I_{n,j}$. Then the expectation of the product $(g - 1_A)(x)(g - 1_A)(x+h)(g - 1_A)(x+k)(g - 1_A)(x+h+k)$ is 0, by independence. Now consider the integral of this product over values of x, h, k such that at least two vertices of the corresponding parallelogram lie in the same interval. This integral is at most the Haar measure of the set D of these parallelograms,

$D = \{(x, h, k) \in \mathbb{T}^3 : \exists \omega, \omega' \in \{0, 1\}^2, \omega \neq \omega', \exists j \in [2^n], I_{n,j} \ni x + \omega \cdot (h, k), x + \omega' \cdot (h, k)\}$, where $\omega \cdot (h, k) = \omega_1 h + \omega_2 k$. Note that D is the support of the measurable function

$$F_D : (x, h, k) \mapsto \sum_{\substack{\omega, \omega' \in \{0, 1\}^2 \\ \omega \neq \omega'}} \sum_{j \in [2^n]} 1_{I_{n,j}}(x + \omega \cdot (h, k)) 1_{I_{n,j}}(x + \omega' \cdot (h, k)),$$

so D is a measurable subset of \mathbb{T}^3 . We want to show that the Haar measure of D vanishes as $n \rightarrow \infty$, and for that it clearly suffices to show that the integral of F_D vanishes. Fix any pair of distinct ω, ω' . Then $\omega - \omega'$ has at least one non-zero coordinate (where the subtraction here is coordinate-wise). For a fixed x , we have $x + \omega \cdot (h, k), x + \omega' \cdot (h, k)$ in the same interval $I_{n,j}$ only if $(\omega - \omega') \cdot (h, k) \in (-2^{-n}, 2^{-n})$. It follows that, for this pair ω, ω' , the integral

$$\int_{\mathbb{T}^3} \sum_{j \in [2^n]} 1_{I_{n,j}}(x + \omega \cdot (h, k)) 1_{I_{n,j}}(x + \omega' \cdot (h, k)) dx dh dk$$

is at most the Haar measure of the slab $S = \{(h, k) \in \mathbb{T}^2 : (\omega - \omega') \cdot (h, k) \in (-2^{-n}, 2^{-n})\}$, which is 2^{-n+1} . Applying this argument to each pair ω, ω' shows that the Haar measure of D vanishes as required. \square

Remark 4.6. A minor modification of the above proof allows one to replace the U^2 norm by the U^d norm for any $d \geq 2$; these norms are useful generalizations of the U^2 norm [12].

Proof of Lemma 4.1. Let G be either \mathbb{T} or \mathbb{Z}_p , where p is large. For a measurable function $f : G \rightarrow [0, 1]$, let A be the set given by Lemma 4.5 applied with initial parameter ϵ . Then by the Cauchy-Schwarz inequality we have $|\int_G 1_A - \int_G f| \leq \|1_A - f\|_{U^2} \leq \epsilon$. Now let L be any G -admissible linear form in at least three variables. Then multilinearity of S_L and Theorem 2.5 imply $|S_L(f) - S_L(A)| \leq t\|f - 1_A\|_{U^2} \leq t\epsilon$. \square

5. REMOVAL LEMMAS

In this section, for $G = \mathbb{T}$ or \mathbb{Z}_p , we are interested in measurable subsets $A \subset G$ of positive measure such that $S_L(A) < \delta$ for some small $\delta > 0$. We shall prove the following result.

Lemma 5.1. *Let $\epsilon > 0$ and let L be a linear form in t variables. There exists $\delta > 0$ such that, for any measurable sets $A_1, A_2, \dots, A_t \subset \mathbb{T}$ satisfying $S_L(A_1, A_2, \dots, A_t) < \delta$, there are measurable sets $E_i \subset \mathbb{T}$ with $\mu(E_i) < \epsilon$ such that $(A_1 \setminus E_1) \times (A_2 \setminus E_2) \times \dots \times (A_t \setminus E_t) \cap \ker L = \emptyset$.*

This result has a well-known analogue for finite abelian groups, an analogue that now has several proofs. One such proof, given in [6], proceeds by turning the problem into one of removing small subgraphs from a certain auxiliary graph, and then applying a removal lemma for graphs. Unfortunately this argument does not seem to extend straightforwardly to the infinitary setting of the group \mathbb{T} . An earlier proof was given by Green [4], who established a Fourier-analytic regularity lemma from which the finitary removal lemma follows. This proof does generalize to the infinitary setting, allowing one to establish a removal lemma for arbitrary compact abelian groups, but checking this is somewhat technical. Instead, we shall prove Lemma 5.1 using a finitary removal lemma as a black box. We first reduce to the special case where L is the linear form $L(x) = x_1 + \dots + x_t$; this form will be denoted by $\mathbf{1}$ throughout this section.

Lemma 5.2. *For any $\epsilon > 0$ and any positive integer t there exists $\delta > 0$ such that the following holds. Let A_1, A_2, \dots, A_t be measurable subsets of \mathbb{T} such that $S_{\mathbf{1}}(A_1, \dots, A_t) < \delta$. Then there exist measurable sets $E_i \subset \mathbb{T}$ such that $(A_1 \setminus E_1) \times \dots \times (A_t \setminus E_t) \cap \ker \mathbf{1} = \emptyset$ and $\mu(E_i) < \epsilon$ for each i .*

When N is prime, the equivalence between the \mathbb{Z}_N -analogues of Lemmas 5.1 and 5.2 follows simply from inverting the dilations. For the circle we can still prove the desired equivalence at the cost of a slight worsening in the dependence between the parameters. To show that Lemma 5.2 implies Lemma 5.1 we use the following.

Lemma 5.3. *Let $L(x) = c_1x_1 + \dots + c_tx_t$ and let $A_1, A_2, \dots, A_t \subset \mathbb{T}$ be measurable. Then*

$$S_L(A_1, \dots, A_t) \leq S_{\mathbf{1}}(c_1A_1, \dots, c_tA_t) \leq |c_1 \cdots c_t| S_L(A_1, \dots, A_t).$$

Proof. The map $\phi : (x_1, \dots, x_t) \rightarrow (c_1x_1, \dots, c_tx_t)$ is a continuous endomorphism on \mathbb{T}^t which restricts to a continuous surjective homomorphism from $\ker L$ to $\ker \mathbf{1}$. In fact, since L is the composition $\mathbf{1} \circ \phi$, we have $\ker L = \phi^{-1} \ker \mathbf{1}$. By Lemma 2.3 we therefore have $\mu_{\mathbf{1}} = \mu_L \circ \phi^{-1}$. It follows that $S_{\mathbf{1}}(A_1, \dots, A_t) = S_L(c_1^{-1}A_1, \dots, c_t^{-1}A_t)$, where c_i^{-1} denotes taking the preimage under dilation by c_i .

Applying this to the sets c_iA_i gives $S_{\mathbf{1}}(c_1A_1, \dots, c_tA_t) = S_L(c_1^{-1}c_1A_1, \dots, c_t^{-1}c_tA_t)$, and it is easy to check that $c_i^{-1}c_iA_i = \bigcup_{j \in \mathbb{Z}_{c_i}} (A_i + j/c_i)$. The claim then follows from basic properties of S_L , namely multilinearity and invariance under translation by elements of $\ker L$. \square

Proof of Lemma 5.1 using Lemma 5.2. Fix $\epsilon > 0$ and let $\delta > 0$ be as given by Lemma 5.2. Now suppose $A_1, A_2, \dots, A_t \subset \mathbb{T}$ are measurable subsets satisfying $S_L(A_1, \dots, A_t) < \delta$

$\delta/|c_1 \cdots c_t|$. By Lemma 5.3 we then have $S_1(c_1 A_1, \dots, c_t A_t) < \delta$. Lemma 5.2 therefore provides us with sets F_i with $\mu(F_i) < \epsilon$ such that $(c_1 A_1 \setminus F_1) \times \cdots \times (c_t A_t \setminus F_t) \cap \ker \mathbf{1} = \emptyset$. Let $E_i = c_i^{-1} F_i$, and let ϕ be the map from Lemma 5.3. Since

$$\phi((A_1 \setminus E_1) \times \cdots \times (A_t \setminus E_t) \cap \ker L) = (c_1 A_1 \setminus F_1) \times \cdots \times (c_t A_t \setminus F_t) \cap \ker \mathbf{1} = \emptyset,$$

the sets E_i satisfy the conclusion of Lemma 5.1. \square

We shall now prove Lemma 5.2. For notational convenience we assume in the proof that $\mathbf{1}$ has last coefficient $c_t = -1$, without loss of generality. Note that as a special case of (2.4) we have

$$S_1(A_1, \dots, A_t) = \int_{\mathbb{T}^{t-1}} 1_{A_1}(x_1) \cdots 1_{A_{t-1}}(x_{t-1}) 1_{A_t}(x_1 + x_2 + \cdots + x_{t-1}) dx_1 \cdots dx_{t-1}. \quad (5.1)$$

Using Lemma 4.3 one may reduce the proof of Lemma 5.2 to establishing the following variant.

Lemma 5.4. *For any $\epsilon > 0$ and $t \in \mathbb{N}$ there exists $\delta > 0$ such that the following holds. Let $A_1, A_2, \dots, A_t \subset \mathbb{T}$ be unions of intervals $[(j-1)/N, j/N)$, $j \in [N]$, for some $N \in \mathbb{N}$. Suppose further that $S_1(A_1, \dots, A_t) < \delta$. Then there are measurable sets $E_i \subset A_i$ with $\mu(E_i) < \epsilon$ such that $(A_1 \setminus E_1) \times \cdots \times (A_t \setminus E_t) \cap \ker \mathbf{1} = \emptyset$.*

Proof of Lemma 5.2 using Lemma 5.4. Fix $\epsilon > 0$ and let $\delta \in (0, \epsilon)$ be such that Lemma 5.4 holds with initial parameter $\epsilon/2$. Let $A_1, \dots, A_t \subset \mathbb{T}$ be measurable sets satisfying $S_1(A_1, \dots, A_t) < \delta/(t+1)$. Using (5.1) and a telescoping expansion (multilinearity of S_1) we have $|S_1(A_1, \dots, A_t) - S_1(B_1, \dots, B_t)| \leq \mu(A_1 \Delta B_1) + \cdots + \mu(A_t \Delta B_t)$ for any measurable sets B_1, \dots, B_t . From Lemma 4.3 we obtain a positive integer n and sets B_i that are unions of intervals $[(j-1)2^{-n}, j2^{-n})$ such that $\mu(A_i \Delta B_i) \leq \delta/(t+1)$ for each $i \in [t]$. It follows that $S_1(B_1, \dots, B_t) < \delta$. We apply Lemma 5.4 to these discretized sets, obtaining measurable sets F_1, \dots, F_t such that $\mu(F_i) < \epsilon/2$ and

$$(B_1 \setminus F_1) \times \cdots \times (B_t \setminus F_t) \cap \ker \mathbf{1} = \emptyset.$$

Now letting $E_i = F_i \cup (A_i \setminus B_i)$, we have $\mu(E_i) < \epsilon$, and $A_i \setminus E_i \subset B_i \setminus F_i$, whence $(A_1 \setminus E_1) \times \cdots \times (A_t \setminus E_t) \cap \ker \mathbf{1} = \emptyset$ as required. \square

We shall deduce Lemma 5.4 from the following analogue for \mathbb{Z}_N , proved in [4].

Lemma 5.5. *For any $\epsilon > 0$ and $t \in \mathbb{N}$ there exists $\delta > 0$ such that the following holds. Let $N \in \mathbb{N}$ and suppose that $A_1, \dots, A_t \subset \mathbb{Z}_N$ and that $S_1(A_1, \dots, A_t) < \delta$. Then there are sets $E_i \subset A_i$ with $|E_i| \leq \epsilon N$ such that $(A_1 \setminus E_1) \times \cdots \times (A_t \setminus E_t) \cap \ker \mathbf{1} = \emptyset$.*

To use this we need to express the solution measure S_1 on \mathbb{T} in terms of solution measures on \mathbb{Z}_N . Recall that the Eulerian number $\langle n \rangle_k$ is the number of permutations a_1, \dots, a_n of $[n]$ in which there are precisely k values of i such that $a_i < a_{i+1}$.

Lemma 5.6. *Let $t \geq 3$ and suppose $A_1, \dots, A_t \subset \mathbb{T}$ are unions of intervals $[(j-1)/N, j/N)$, $j \in [N]$, for some positive integer N . Let $A'_i \subset \mathbb{Z}_N$ be defined by $1_{A'_i}(x) = 1_{A_i}(x/N)$. Then*

$$S_1(A_1, \dots, A_t) = \frac{1}{(t-1)!} \sum_{r=0}^{t-2} \left\langle t-1 \right\rangle_r S_1(A'_1, \dots, A'_t - r).$$

In particular, $S_1(A'_1, \dots, A'_t - r) \leq (t-1)! S_1(A_1, \dots, A_t)$ for any $r \in \{0, 1, \dots, t-2\}$.

Proof. For any $x \in \mathbb{T}$ we have $1_{A_i}(x) = 1_{A_i}(\lfloor Nx \rfloor / N)$, where multiplication by N , floor function, and division by N are all taken in \mathbb{R} . Then $S_1(A_1, \dots, A_t)$ equals

$$\begin{aligned} & \int_{\mathbb{T}^{t-1}} 1_{A_1}(\lfloor Nx_1 \rfloor / N) \cdots 1_{A_t}(\lfloor N(x_1 + \cdots + x_{t-1}) \rfloor / N) dx_1 \cdots dx_{t-1} \\ &= \sum_{a_1, \dots, a_{t-1} \in \mathbb{Z}_N} 1_{A_1}(a_1/N) \cdots 1_{A_{t-1}}(a_{t-1}/N) \\ & \quad \cdot \int_{[0, 1/N]^{t-1}} 1_{A_t}(\lfloor N((a_1/N + y_1) + \cdots + (a_{t-1}/N + y_{t-1})) \rfloor / N) dy_1 \cdots dy_{t-1} \\ &= \frac{1}{N^{t-1}} \sum_{a_1, \dots, a_{t-1} \in \mathbb{Z}_N} 1_{A'_1}(a_1) \cdots 1_{A'_{t-1}}(a_{t-1}) \\ & \quad \cdot \int_{[0, 1]^{t-1}} 1_{A'_t}(a_1 + \cdots + a_{t-1} + \lfloor y_1 + \cdots + y_{t-1} \rfloor) dy_1 \cdots dy_{t-1}. \end{aligned}$$

For any real-valued function f on the integers, [3, 6.65] gives

$$\int_0^1 \cdots \int_0^1 f(\lfloor x_1 + \cdots + x_n \rfloor) dx_1 \cdots dx_n = \sum_{k=0}^{n-1} \left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle \frac{f(k)}{n!}.$$

The result now follows immediately. \square

Proof of Lemma 5.4. We can assume that $t \geq 3$. Given $\epsilon > 0$, set $\epsilon' = \epsilon/(t-1)$ and let δ' be given by Lemma 5.5 applied with initial parameter ϵ' . Define $\delta = \delta'/(t-1)!$ and let $A_1, \dots, A_t \subset \mathbb{T}$ be sets that are unions of intervals of the form $[j/N, (j+1)/N)$ for some positive integer N and satisfy $S_1(A_1, \dots, A_t) < \delta$.

Let the sets $A'_i \subset \mathbb{Z}_N$ be as in Lemma 5.6. We then have $S_1(A'_1, \dots, A'_t - r) < \delta'$ for each $r = 0, \dots, t-2$, and so Lemma 5.5 gives us sets $E'_{i,r} \subset A'_i$ such that $|E'_{i,r}| \leq \epsilon'N$ and

$$(A'_1 \setminus E'_{1,r}) \times \cdots \times ((A'_t - r) \setminus (E'_{t,r} - r)) \cap \ker \mathbf{1} = \emptyset. \quad (5.2)$$

Setting $E'_i = \bigcup_r E'_{i,r}$ we therefore see that there are no solutions to $y_1 + \cdots + y_{t-1} = y_t - r$ with $y_i \in A'_i \setminus E'_i$ for $r \in [0, t-2] \subset \mathbb{Z}_N$.

We now define corresponding removal-sets in \mathbb{T} : for each $i \in [t]$ let

$$E_i = \bigcup_{x \in E'_i} [x/N, (x+1)/N).$$

Since $|E'_i| \leq \epsilon N$ we have $\mu(E_i) \leq \epsilon$ for each i . Now suppose that we had an element

$$(x_1, \dots, x_t) \in (A_1 \setminus E_1) \times \cdots \times (A_t \setminus E_t) \cap \ker \mathbf{1}.$$

Then, since each of the sets $A_i \setminus E_i$ is a union of intervals $[j/N, (j+1)/N)$, we have that the element $\lfloor Nx_i \rfloor$ of \mathbb{Z}_N lies in $A'_i \setminus E'_i$. But we also have $x_t = x_1 + \cdots + x_{t-1}$, whence $\lfloor Nx_t \rfloor = \lfloor Nx_1 \rfloor + \cdots + \lfloor Nx_{t-1} \rfloor + r$ for some $r \in [0, t-2] \subset \mathbb{Z}_N$. This contradicts (5.2), and therefore $(A_1 \setminus E_1) \times \cdots \times (A_t \setminus E_t) \cap \ker \mathbf{1} = \emptyset$ as required. \square

6. PROOFS OF CONVERGENCE AND POSITIVITY

We can now prove our main result.

Proof of Theorem 1.3. We are given a finite family \mathcal{F} of linear forms. Let n be the cardinality of \mathcal{F} , and fix $\epsilon > 0$. Let $\delta \in (0, \epsilon/2)$ be such that Lemma 5.1 and its analogue for \mathbb{Z}_p both work for each $L \in \mathcal{F}$, with initial parameter $\epsilon/2n$. Let t be the maximum number of variables that occurs in a form in \mathcal{F} and let $C = C(\epsilon, \mathcal{F})$ be such that Corollary 3.8 (transference) and Lemma 4.1 (functions to sets) both work when applied with initial parameter $\delta/2t$ for any $p > C$ and $L \in \mathcal{F}$. We claim that $|d_{\mathcal{F}}(\mathbb{T}) - d_{\mathcal{F}}(\mathbb{Z}_p)| \leq \epsilon$ for any prime $p > C$.

To see that $d_{\mathcal{F}}(\mathbb{T}) \geq d_{\mathcal{F}}(\mathbb{Z}_p) - \epsilon$, let $\alpha = d_{\mathcal{F}}(\mathbb{Z}_p)$ and let $A \subset \mathbb{Z}_p$ be \mathcal{F} -free with $\mu(A) = \alpha$. Corollary 3.8 then gives us a measurable function $g : \mathbb{T} \rightarrow [0, 1]$ with $\int_{\mathbb{T}} g = \alpha$ such that $S_L(g) < \delta/2$ for every $L \in \mathcal{F}$. Applying Lemma 4.1 to g with initial parameter $\delta/2t$, we obtain a measurable subset B of \mathbb{T} of density at least $\alpha - \delta$ such that $S_L(B) < \delta$ for every $L \in \mathcal{F}$.

We now apply the removal lemma on \mathbb{T} . By our choice of δ , Lemma 5.1 gives us an \mathcal{F} -free subset D of B with $\mu(D) \geq \mu(B) - n\epsilon/2n \geq \alpha - \delta - \epsilon/2$. Therefore $d_{\mathcal{F}}(\mathbb{T}) \geq d_{\mathcal{F}}(\mathbb{Z}_p) - \epsilon$ as required.

The same argument, but with the roles of \mathbb{Z}_p and \mathbb{T} swapped, shows that we also have $d_{\mathcal{F}}(\mathbb{Z}_p) \geq d_{\mathcal{F}}(\mathbb{T}) - \epsilon$, and so the result follows. \square

Remark 6.1. One of our aims for the argument above was to treat the direction $d_{\mathcal{F}}(\mathbb{Z}_p) > d_{\mathcal{F}}(\mathbb{T}) - \epsilon$ and its opposite in a unified manner. It should be noted however that the first direction can also be treated more directly, without using Fourier analysis. In a nutshell, if $A \subset \mathbb{T}$ is \mathcal{F} -free with $\mu(A) > d_{\mathcal{F}}(\mathbb{T}) - \epsilon$ then the continuity of the forms in \mathcal{F} implies that one can find an \mathcal{F} -free open set A' of measure at least $d_{\mathcal{F}}(\mathbb{T}) - 2\epsilon$, and then for large p the set $A_p = \{x \in \mathbb{Z}_p : x/p \in A'\}$ is \mathcal{F} -free and of density at least $d_{\mathcal{F}}(\mathbb{T}) - 3\epsilon$ in \mathbb{Z}_p .

We now prove Proposition 1.4, which said that $d_{\mathcal{F}}(\mathbb{T})$ is positive for any finite family \mathcal{F} of non-invariant forms. We can do this by modifying an idea employed by Ruzsa [9]. For a form $L(x) = c_1x_1 + \dots + c_tx_t$ we write $s_L = |c_1| + \dots + |c_t|$.

Proposition 6.2. *Let \mathcal{F} be a finite family of non-invariant linear forms. Then*

$$d_{\mathcal{F}}(\mathbb{T}) \geq \left(\sum_{L \in \mathcal{F}} s_L \right)^{-1}.$$

Proof. Let $s = \sum_{L \in \mathcal{F}} s_L$ and let A be the interval $(-1/2s, 1/2s)$ embedded in \mathbb{T} . We claim that there is a translate $A - y$ of A that is \mathcal{F} -free. This will be the case if $L(y, \dots, y) \notin L(A \times \dots \times A) = (-s_L/2s, s_L/2s)$ for each $L \in \mathcal{F}$. Since each L is non-invariant, each such condition on y excludes a finite union of open intervals, with total length s_L/s . Thus all the conditions together exclude a finite union of open intervals, the lengths of which sum to at most 1. Hence there is some $y \in \mathbb{T}$ outside this union of intervals, and the result follows. \square

7. CONCLUDING REMARKS

Our proof of Theorem 1.3 consists essentially of a combination of the transference result Proposition 3.4 with the removal result Lemma 5.1. Two remarks should be made about this. The first is that the combination of Fourier-regularization (a key tool in the proof of the transference result) and removal results has been used successfully before; in particular, Green [4, Theorem 9.3] used it to relate the number of subsets of $[N]$ that are free from non-trivial solutions to $L(x) = 0$ to the maximum size of an L -free subset of $[N]$. The second remark is that the argument in our proof of Theorem 1.3 readily yields analogues of the theorem for many other families of groups, provided in particular that the appropriate removal lemmas are available (examples of such families include (\mathbb{Z}_{kp}) , where $k \in \mathbb{N}$ is fixed and p ranges over the primes). We have not treated such generalizations here in order to avoid certain technicalities. Let us note, however, that if all one is interested in is convergence of maximal densities, rather than convergence to a particular quantity on a group, then all the theory one needs is finitary and the appropriate removal lemmas are well-known.

It would be interesting to generalize Theorem 1.3 to allow the family \mathcal{F} to consist not just of single linear equations but also of systems of linear equations. To this end it can be useful to classify systems according to a notion of *complexity* related to the Gowers norms (see [2]). The methods of this paper readily extend to give convergence of $d_{\mathcal{F}}(\mathbb{Z}_p)$ for a family of systems of complexity 1, but establishing \mathbb{T} as a limit group along these lines requires an extension of Lemma 5.1. Convergence for systems of greater complexity requires other methods.

Finally, regarding the original question of Ruzsa mentioned in the introduction, we note that there is a simple transference result for functions on $[N]$, proved using an argument somewhat different from that employed for transference here, and that this can be used to answer Ruzsa's question affirmatively (we shall detail this elsewhere).

Acknowledgements. The authors are grateful to Tom Sanders for helpful comments and to an anonymous referee for pointing out the argument in Remark 6.1. The second-named author would also like to thank Ben Green for his supervision and encouragement that led to the main results of [10, Chapter 4].

REFERENCES

- [1] E. Croot, The minimal number of three-term arithmetic progressions modulo a prime converges to a limit, *Canad. Math. Bull.*, **51** (2008), 47–56.
- [2] W. T. Gowers and J. Wolf, The true complexity of a system of linear equations, *Proc. Lond. Math. Soc.*, (3) **100** (2010), no. 1, 155–176.
- [3] R. L. Graham, D. E. Knuth and O. Patashnik, *Concrete Mathematics*, Addison-Wesley (1989).
- [4] B. J. Green, A Szemerédi-type regularity lemma in abelian groups, with applications, *Geom. Funct. Anal.* **15** (2005), no. 2, 340–376.
- [5] B. Green and T. Tao, Linear equations in primes, *Annals of Math.* **171** (2010), no. 3, 1753–1850.
- [6] D. Král', O. Serra and L. Vena, A combinatorial proof of the removal lemma for groups, *J. Combin. Theory Ser. A* **116** (2009), no. 4, 971–978.
- [7] W. Rudin, *Fourier analysis on groups*, Interscience Publishers (1962).
- [8] I. Z. Ruzsa, Solving a linear equation in a set of integers, I, *Acta Arith.* **65** (1993), 259–282.
- [9] I. Z. Ruzsa, Solving a linear equation in a set of integers, II, *Acta Arith.* **72** (1995), 385–397.

- [10] O. Sisask, *Combinatorial properties of large subsets of abelian groups*, Ph.D. Thesis, University of Bristol (2009).
- [11] T. Tao, *Structure and Randomness: pages from year one of a mathematical blog*, American Mathematical Society (2008).
- [12] T. Tao and V. Vu, *Additive combinatorics*, Cambridge University Press (2006).

CENTRE FOR MATHEMATICAL SCIENCES, WILBERFORCE ROAD, CAMBRIDGE CB3 0WB, UNITED KINGDOM

E-mail address: `pc308@cam.ac.uk`

SCHOOL OF MATHEMATICAL SCIENCES, QUEEN MARY, UNIVERSITY OF LONDON, MILE END ROAD, LONDON E1 4NS, UNITED KINGDOM

E-mail address: `O.Sisask@qmul.ac.uk`